DOCUMENT RESUME

ED 211 499 SP 019 332

AUTHOR Lawrence, Gordon
TITLE A Synthesis of Research on the Effectiveness of Staff Development Programs.
PUB DATE May 81
NOTE 13p.; Paper presented at the State of Ohio Department of Education Annual Conference on Inservice Education (Columbus, OH, May 5, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Change Agents; *Educational Research; *Inservice Teacher Education; Participant Satisfaction; Postsecondary Education; *Program Effectiveness; Program Evaluation; *Research Methodology; *Staff Development; Teacher Attitudes; Teacher Behavior

ABSTRACT

A report is given of a comprehensive review of research on inservice education. The review process began with 6,000 abstracts and references. About 150 of the documents reported some quantitative findings. Fifty-nine studies met topic criteria and also had statistics which could be synthesized. These studies were sorted into four categories of objectives: cognitive, affective, performance, and consequence. Cognitive objectives dealt with the participant's store of information or concepts, affective objectives concerned changes in attitudes. Performance objectives included observable skills and behaviors of participants, and consequence objectives were those that concerned changes in pupils resulting from what participants gained in the program. The review was directed toward finding inservice program contexts, materials, and procedures that are most effective in facilitating cognitive, affective, performance, or consequence change in participants. Eleven general findings are reported on the inservice programs most successful in accomplishing their objectives. A description is given of three successful inservice programs that were well designed and that tended to treat teachers as professionals. The programs also illustrate three different research designs that are manageable in conducting inservice programs. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED211499

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Gordon Lawrence
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A SYNTHESIS OF RESEARCH ON THE EFFECTIVENESS OF STAFF DEVELOPMENT PROGRAMS

Gordon Lawrence
University of Florida

I imagine I speak for most of us here when I say that inservice education for teachers has an image problem. Most teachers I have worked with have held in their minds an image of inservice education that does not bode well for the success of inservice programs. Their mental image is what they imagine will happen in the name of inservice education. See if the following teacher's account sounds familiar to you.

Some hot Wednesday afternoon in the spring we gather in the media center and somebody puts in my hand a three-page set of questions with little boxes, and I'm supposed to check the boxes that match my needs for inservice education. Then in the fall I'm handed an inservice schedule sheet that lists workshop topics and the names of some university professors and a few others who have been engaged to conduct the workshops on certain dates. I don't recognize any topics related to the little boxes I checked last spring, but I'm not surprised.

Because I'm expected to participate in at least some inservice activities, I sign up for two or three. When I get to them, I am given the general impression that I am deficient in something that researchers or theorists regard as important, so I should be remediated. The leader, noticing that I am less than enthusiastic about this workshop, makes me feel vaguely guilty in not taking more initiative to shape the workshop to my needs. So here I am after a hard day's work, low on energy, feeling guilty because I'm giving this outsider a poor impression of our faculty, feeling angry about feeling guilty, feeling sorry for having signed up, feeling really angry at an inservice system that gets me into situations like this. It does not call upon the

Paper presented at the State of Ohio Department of Education Annual Conference on Inservice Education (Columbus, OH, May5, 1981).

SP 019 332

best in me. It makes me feel uncreative and less that professional, and I don't know why.

Does this account ring true? I am afraid that such experiences have shaped the image that most teachers have of inservice education. We all know about the power of expectations, the power of beliefs in moulding our behavior. How can expectations for inservice education be positive when its image is negative?

If I say that we at this conference have the job of reshaping the image of inservice education, I am not suggesting a Madison Avenue public relations and advertising campaign. Nor designing new packaging for the same familiar product. Nor cosmetic changes like a new name--whether we call it staff development, or professional development or something else, the mental image that teachers have of inservice education depends on what they experience in its name.

If we are to take steps to reshape the collective mental image that teachers have of inservice education, we must first consider our own images. My personal image of what inservice education can be--what it can realistically become- has come largely from my examination of the research on inservice programs and on why some programs are more successful than others. My task here is to share with you the generalizations that can be drawn from the research and used to guide policy plans for inservice education.

Seven years ago, some graduate students and I undertook the first comprehensive search and review of research on inservice education. Among 800 reports on inservice programs that we located, we identified 97 as being research in the sense that they gave some attention to matters of reliability and validity of the documentation process. Just this past year we undertook an update of the earlier review and included studies that

appeared since 1974. This time we wanted to use a new technique called meta analysis to numerically combine effects across studies.

Meta Analysis

In meta analysis, any gains or changes in effects of a program can be converted to a common base and combined across independent studies. For example, suppose two studies examined teacher classroom performance of certain skills as outcomes of staff development activities. Assume that one was concerned with reading instruction skills and the other with skills of questioning. If the researchers were looking at different skills and had used different observation instruments, how could the effects in the two studies be combined? By regarding the two sets of outcomes as being of a larger class--teacher performance of skills--we can say that these are both performance measures. By standardizing the scores, characterizing each as the size of the effect in relation to standard deviation, the two sets of scores can be put on the same scale.

The review process began with 6,000 abstracts and references. About 150 of the documents reported some quantitative findings. Those that met the topical criteria and also had statistics which could be synthesized across studies numbered just 59.

All findings of the 59 studies were sorted into four categories of objectives (dependent variables): cognitive, affective, performance and consequence. Cognitive objectives were those dealing with the participant's store of information or concepts, typically measured by paper and pencil tests. Affective objectives concerned changes in attitudes, values and interests of participants, often measured by self-report instruments. Performance objectives included observable skills and behaviors of participants, typically reported by trained observers from live or taped events. Consequence

objectives were those that concerned changes in pupils resulting from what participants gained in the program. Consequence outcomes may be pupil achievement, attitude or behavior changes.

Thus, the review was directed toward answering four questions. When the purpose of professional development programs is cognitive change in the participants, what contexts, materials and procedures have been found to be more effective for that purpose? The other three questions similarly dealt with affective, performance and consequence changes.

Findings of the Review

There were many detailed findings that I could not cover in these few minutes. I will discuss 11 general findings.

The inservice programs most successful in accomplishing their objectives were ones that:

-involved teachers actively in initiating, planning and conducting the program;

-were designed as a collective effort of a faculty, with common purposes and directed toward general faculty development rather than focusing on the separate goals and needs of individual faculty members;

-were funded in ways that permitted the teachers and administrators of individual schools to sponsor them, to design activities, and to select inside and outside leadership as appropriate to the plans (Programs led by school supervisors, teachers, and college personnel met their objectives better than those led by state department of education personnel, school district staff, or other outside consultants);

-were scheduled at times (evenings, summers) that did not

compete with but complemented other professional obligations of the participants (Programs scheduled during work hours were considerably less successful in achieving objectives.);

-had diverse program patterns that seemed to emphasize teacher responsibility--self-instruction, peer study groups, college courses, one-to-one consultation (The formats loosely labeled as workshop and staff meeting were considerably less successful in achieving objectives.);

-involved participants in both receptive and active roles--receiving new ideas and putting them into action--rather than one type without the other;

-had sequences in which participants could try out new things in their classrooms (or in simulations) and then receive appropriate feedback from a skilled person (Programs in which participants were expected to store up new ideas and behavior prescriptions for a future time were distinctly less successful in achieving objectives.);

-had leaders who were linkers with a university or other center concerned with professional development;

-had opportunities for participants to see demonstrations of exemplary practices, and to learn the skills of observing the practices in themselves and others;

-did not rely on lecture presentations as the main activity;

-were conducted at the school site if the programs emphasized affective or skill performance objectives.

One of the variables that did not account for differences in outcomes was whether or not teachers had a choice of participating in the program.

A reader of the 59 studies gets the distinct impression that the more successful programs treated their participants as professionals, as conscientious people who want to continue to expand their skills and their range of professional competence. The more successful programs seemed not only to give careful attention to planning the inservice activities, but were planned to fit the design and methods to the people--to their purposes and to the context in which they worked.

Some of you may be familiar with the four-year Rand study of 293 federal seed money projects. While it did not produce quantitative data that could be included in our meta analysis, it did produce generalizations that were strikingly similar to those of our study. The Rand findings highlighted the importance of total faculty involvement in staff development--a developmental model rather than a deficit model. Successful programs had solid institutional support. In the activities themselves, teachers had active roles in continued planning, in identifying what their training should be, in local materials development, in making classroom applications, and in evaluating outcomes.

For the second half of my talk, I want to describe to you three successful inservice programs that turned up in the studies we reviewed. I believe they represent programs that were well designed to treat teachers as professionals. They also illustrate three different research designs that are manageable in the conducting of inservice programs.

Four Successful Studies and Their Research Designs

A program of self-instruction in classroom skills. Anderson, Evertson and Brophy (1979) trained observers to visit classrooms of 20 first grade

teachers, 10 of whom had participated in a brief training session on techniques for conducting reading groups. Seven other teachers attended the training and agreed to try to implement the techniques but were not observed. Observations were made over a six-month period to examine the actual extent of implementation of the techniques.

The professional development activity was simple, brief and inexpensive. The researchers met with teachers, explained the plan, and gave those who were interested a brief manual describing 22 principles of managing reading instruction drawn from research findings. Those who read the material and wanted to continue met one more time with the researchers for questions and answers about the 22 principles. This simple program, and the manual itself, treated the teachers as professionals who would be motivated to adapt relevant research findings to their classroom practices if they saw the findings as serving students' needs. The program depended upon teacher initiative and self-instruction.

Students' reading scores at the beginning and end of the school year provided a consequence measure of the effects of the intervention. The students of participant teachers, both the observed and unobserved groups, scored significantly higher than students of the control teachers, about one standard deviation higher (ES = .90). Observers' records also showed substantial implementation of the 22 principles by participants, with the difference between treatment and control groups showing a mean effect size of .72. Linear regression models were used to analyze reading scores and analyses of variance were computed for observation data.

A Workshop Based on Modeling and Simulation. Guttinger and Hines (1977) evaluated a set of workshops designed for diffusing a laboratory approach to individualized, developmental reading instruction. Reading laboratory directors, principals and reading teachers of

three middle schools and one high school participated in a sequence of workshops in which they went through the experience of being students in the reading laboratory program -- testing, interpretation of results, counseling and goal setting. This simulation was followed by analyzing case studies of pupils, role planning and practice in using and managing the materials, equipment and procedures of the laboratory. Exemplary practices in running a reading laboratory were modeled, with participants then practicing what was modeled, with feedback from workshop leaders.

After a four-day orientation workshop, participants helped to schedule and plan a three-or four-day workshop at their own schools. Follow up consultation help was available after that. Each school that adopted the laboratory model had a subsequent two-day workshop for content-area teachers. Availability of follow-up support was regarded as an important element of the program model, as was the sharing of leadership responsibilities with participants.

The dependent measure was the subsequent gain in reading scores of pupils who used the labs in the four schools. Complete data sets were obtained on 1014 students. An experimental, pre-post, control group design was used, with students randomly assigned to classrooms and to treatments. Treatment students' gains significantly exceeded those of students who had regular reading instruction. The difference was nearly one standard deviation (ES = .97).

Practice with case examples using tapes and films. Kasdon and Kelly (1969) used films and audio tapes to train teachers in the process of identifying pupils' reading instruction levels. In this simulated setting, the viewers administered an informal reading inventory to pupils on film. They also were taught techniques of analyzing the reading level of various instructional materials. Comparing, analyzing and discussing protocols with

fellow participants was regarded as an important feature of the program. The treatment encompased five two-hour sessions.

Three stratified random samples of 32 teachers each from grades 2 through 5 were involved in the three-group, posttest-only, control group design. One treatment group received instruction in August before school, one in October after classroom reading groups had been formed. The dependent variable was the size of the gap between the reading level of texts that teachers assigned to pupils and the pupils' scores on a reading inventory. One pupil was selected randomly from the classroom of each teacher in the three groups, for a total of 96 gap scores. Both treatment groups had gap scores significantly lower than the control, but the August group showed the greater difference (ES = .71 for August group; ES = .24 for October group). Teachers who received the instruction prior to meeting pupils and assigning them to reading groups showed the greater ability to match reading material to the student's reading level. Timing in the school calendar appears to be an important context variable not generally considered in the design of research on professional development programs.

Observations on the State of the Art

The three examples of well-documented programs have been described here to make the point that solid experimental and qual-experimental research can be accomplished with professional development programs--an area often regarded as too "messy" for research activity. Moreover, the responsibility for research can and should be carried more by school systems. Two-thirds of the 59 studies were initiated by university personnel, including doctoral students undertaking dissertations. The fact prompts the question of why school systems do not pursue program evaluations and research that enters the pool of research literature. Tight budgets may be one reason. Those who shape policy for

schools and decide on budgets have the responsibility of asking whether the pay off from evaluation and research will warrant the investment. If the hesitancy of policy makers is due in part to a lack of good models of research plans suitable for staff development programs, the lack has been remedied.

The value of documentation in professional development programs seems to be little recognized. Only eight of the studies seemed to be part of a general documentation plan. Until documentation becomes a regular part of the work of those who manage the programs, it is difficult to see how policy and budget makers can have much confidence that what works well once can be counted on to be effective from one group to the next, one year to the next, one setting to another, one set of leaders to another.

While the state of the art of program documentation seems primitive, several encouraging impressions emerged from the review, impressions of positive trends. Reporting on programs seems to be improving, becoming more thorough and more comprehensive. Reports in the 1960's and early 1970's commonly gave the topic of a program, then described the research procedures and outcomes, and outcomes, and often neglected specifics about how the program was organized and conducted. The reviewing team had difficulty in classifying many studies into the 25 categories because of missing information.

Program evaluation generally seems to be better planned in the more recent reports. In many NDEA institutes and other programs that drew upon federal funds, the evaluation seemed to be tacked on for perfunctory bureaucratic compliance rather than for genuine interest in evaluation and communication within the professional community. Most of the recent reports show signs that evaluation was build into the program plan at the beginning.

As classifications and concepts such as those used in this review

become refined into a common professional language, the prospects brighten for improved evaluation and for better communication of research findings.

References

Anderson, L. A., Evertson, C. A., & Brophy, J. E. An experimental study of effective teaching in first-grade reading groups. Elementary School Journal, 1979, 79(4), 193-224.

Glass, G. V. Intergrating findings: The meta-analysis of research. Review of Research in Education, 1977, 5, 351-379.

Guttinger, H. I., & Hines, V. A. Field testing and diffusion of an experiment in developmental, individualized reading at the middle and high school levels. Gainesville, Florida: P. K. Yonge Laboratory School, University of Florida, 1977.

Kasdon, L. M., & Kelly D. Simulation: In-service education for teachers of reading. Journal of Experimental Education, 1969, 38(1), 79-86.

Lawrence, G., Baker, D., Hansen, B. & Elzie, P. Patterns of Effective Inservice Education. Florida Department of Education, 1974 (Available from PAEC, Post Office Drawer 190, Chipley, FL 32428).

Lawrence, G. & Harrison, D. Policy Implications of the Research on the Professional Development of Education Personnel: An Analysis of Fifty-Nine Studies. 1981 Report on Educational Personnel Development. Washington, D.C.: Feistritzer Publications, 1980.